

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2002, 14:14:28 : Search time 30.15 Seconds
(without alignments)
1663.964 Million cell updates/sec

Title: US-09-808-885-1
Perfect score: 1467
Sequence: 1 MKRNSQKIKKRRRAKQ.....RQKREERIKRLEESKRP 290

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Sealed: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_19:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	855	58.3	234	11 P70488	P70488 rattus norv
2	792	54.0	655	13 Q9W604	Q9W604 gallus gall
3	327	22.3	513	10 Q9LW95	Q9LW95 nicotiana t
4	317.5	21.6	932	5 Q9VZP5	Q9VZP5 drosophila
5	315	21.5	1144	5 Q9NJH7	Q9NJH7 drosophila
6	314	21.4	390	5 Q18401	Q18401 caenorhabdi
7	305	20.8	522	10 Q9FX85	Q9FX85 arabidopsis
8	284.5	19.4	1224	10 Q9SR21	Q9SR21 arabidopsis
9	283.5	19.3	2081	10 Q9LH98	Q9LH98 arabidopsis
10	283	19.3	363	10 Q93Y04	Q93Y04 arabidopsis
11	281	19.2	679	5 Q95XW8	Q95XW8 caenorhabdi
12	279.5	19.1	260	10 Q9LG29	Q9LG29 arabidopsis
13	276.5	18.8	1280	10 Q9SRD2	Q9SRD2 arabidopsis
14	272	18.5	1166	10 Q9SYF6	Q9SYF6 arabidopsis
15	267.5	18.2	791	13 Q9DGL1	Q9DGL1 fugu rubrip
16	266	18.1	1192	5 Q96127	Q96127 plasmodium

17	265	18.1	695	5 Q9BHM3	Q9BHM3 paramecium
18	262	17.9	368	5 Q45198	Q45198 caenorhabdi
19	259	17.7	312	5 P91570	P91570 caenorhabdi
20	258	17.6	1661	5 Q06166	Q06166 plasmodium
21	257.5	17.6	1146	10 Q9SRD1	Q9SRD1 arabidopsis
22	256.5	17.5	2274	5 Q9VYU0	Q9VYU0 drosophila
23	253	17.2	673	5 Q9U0N1	Q9U0N1 plasmodium
24	251.5	17.1	806	5 Q27529	Q27529 caenorhabdi
25	249.5	17.0	665	5 Q96229	Q96229 plasmodium
26	248.5	16.9	554	5 Q9VPS3	Q9VPS3 drosophila
27	248	16.9	1190	5 Q62235	Q62235 caenorhabdi
28	247.5	16.9	806	5 Q17966	Q17966 caenorhabdi
29	245.5	16.7	335	5 Q76719	Q76719 caenorhabdi
30	244.5	16.7	792	12 Q9YTL7	Q9YTL7 ateline her
31	244	16.6	682	5 Q9GTW3	Q9GTW3 plasmodium
32	242.5	16.5	2083	5 Q9N435	Q9N435 caenorhabdi
33	237	16.2	699	10 Q9ZU59	Q9ZU59 arabidopsis
34	236	16.1	3484	5 P91257	P91257 caenorhabdi
35	235.5	16.1	1108	5 Q9NDI0	Q9NDI0 babesia big
36	235	16.0	1701	5 Q61164	Q61164 plasmodium
37	234.5	16.0	1173	5 Q95XK7	Q95XK7 caenorhabdi
38	234.5	16.0	1812	5 Q9VZD9	Q9VZD9 drosophila
39	232	15.8	1339	11 Q35788	Q35788 rattus norv
40	230	15.7	460	2 Q9ZIU2	Q9ZIU2 borrelia bu
41	227.5	15.5	1336	5 Q22944	Q22944 caenorhabdi
42	226.5	15.4	451	16 Q50870	Q50870 borrelia bu
43	224.5	15.3	385	5 Q17909	Q17909 caenorhabdi
44	224.5	15.3	1062	5 Q960C4	Q960C4 drosophila
45	224	15.3	463	5 Q9VB74	Q9VB74 drosophila

ALIGNMENTS

RESULT 1

ID	P70488	PRELIMINARY:	PRT:	234 AA.
AC	P70488;			
DT	01-FEB-1997 (TREMBLrel. 02, Created)			
DT	01-FEB-1997 (TREMBLrel. 02, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	ANNEXIN V-BINDING PROTEIN (ABP-7) (FRAGMENT).			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-WISTAR; TISSUP-BRAIN;			
RX	MEDLINE=96301899; PubMed=8667030;			
RA	Ohsawa K., Imai Y., Ito D., Kohsaka S.;			
RT	"Molecular cloning and characterization of annexin V-binding proteins with highly hydrophilic peptide structure."			
RL	J. Neurochem. 67:89-97(1996).			
DR	EMBL; D64061; BAA10937.1; -;			
FT	NON_TER 1			
FT	NON_TER 234			
SQ	SEQUENCE 234 AA; 26494 MW; 01BF5910ABE587F5 CRC64;			

Query Match 58.3%; Score 855; DB 11; Length 234;
Best Local Similarity 74.2%; Pred. No. 3.4e-39;
Matches 175; Conservative 23; Mismatches 34; Indels 4; Gaps 3;

QY	15	RRAKKOSFDNDSEEELEDKDSKKTAKPKVE--MYSGSDDDDFNKLPPKAKGAKQSKN 73	
Db	2	QKQKTSFDENSEELEDKDSKPKPNSEVLLSGSEADDPNKLKSKGK-KAQKSTK 60	
QY	74	KWDGSEEDNKKIKERSINSSGSGSEDEFLOSRCQKKNQKPCPNIESGNEDD 133	
Db	61	KRDGSEEDNKKRSKRSRVNSSGSGSEDEFLOSRCQKKNQKSVPTIDSGNEDD 120	
QY	134	DASFKIKTVAQKAEKKERKKRDEKAKLRKLKEELETGKDKQSKQESQKRFEE 193	

```
Db 121 DSSFKTKTVAQKAEKRRERKREBEKAKLRVKEKELEKGRKQSKOREPOKRPDEE 180
QY 194 TVASKVTVDGVPASEEKAETPAEDDNEGDGKKKKKKKKKKKKKKKKKKKKKK 249
Db 181 VLVLRTGTPDAGA--ASEEKGDIATLEDDNEGDGKKKKKKKKKKKKKKKKKKKK 234

RESULT 2
Q9W604 PRELIMINARY; PRT; 655 AA.
AC Q9W604;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE OSTEOBLAST TRANSLATION FACTOR 3F (INITIATION OF TRANSLATION FACTOR 2)
DE (FRAGMENT).
GN Cif2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Weber J.A., Stains J.P., Gay C.V.;
RT "Identification of novel cDNAs from cultured chick osteoblasts by
RT similarity to the sodium-calcium exchanger.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Weber J.A., Gay C.V.;
RT "Cloning of novel cDNAs from chicken osteoblasts by homology to the
RT sodium-calcium exchanger.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF14637; AAD31321.1; -.
DR EMBL; AJ277602; CAB89284.1; -.
FT NON_TER 1
FT NON_TER 655
SQ SEQUENCE 655 AA; 74483 MW; DC33D9B4CDDEC5CF CRC64;

Query Match 54.0%; Score 792; DB 13; Length 655;
Best Local Similarity 44.1%; Pred. No. 1.9e-35;
Matches 193; Conservative 24; Mismatches 55; Indels 166; Gaps 6;

QY 10 KKRDRRAKQSFDDNDSELEDDKSKKTAKPKVEMSGSDDDDFNKLPKKAKGKAQ 69
Db 211 KKRKSKKANLENDYDSEEMEDKDRSKKTQAKQDVJSGS-DDDDLETPKKNKKTQ 269

QY 70 KSNKWDGSEEDNSKKIKERSINSGSGDSEDFLQSRGKQKNQKNKPGPNIESG 129
Db 270 KSNKHLS-EDENYKSKERVGTLSGSGDSEDFSPRGQKNQKPKSTAALGSG 328

QY 130 NEDDASFKITVAQKAEKRRERKRDDEKAKLRKLEKEBELETGCKDQSKQESQRK 189
Db 329 DEEESSEFKVTVAQKAEKRRERKREBEKAKLRKLEKEBELE-GGREAAKPKKAPK 387

QY 190 FEETVSKVTVDGVPASEEK-----AETPAE--- 220
Db 388 AEE-----KASPDVAAPALGEKEKFLQEQRCVNCVSKTVISHDRFYQEARVLTFEGDL 442

QY 221 ----- 220
Db 443 LLYSKFGVVVITWLTSSLLAHMLFFLVCSYSGTVNKSQMVNSHYHYVAVCKHNFSQ 502

QY 221 ----- 220
Db 503 LLLHICVYKSSGPFVTSVTLFFFIKCLYSSLENVPHLELVKCNHQYLELLRFLGS 562

QY 221 -----DDNEGDGKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 271
Db 563 LILSLVADNEDGKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 622
```

```
QY 272 OKREERERIKRLEELSK 289
Db 623 AKREERERIRLEELAK 640

RESULT 3
Q9LW95 PRELIMINARY; PRT; 513 AA.
AC Q9LW95;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE KED.
GN KED.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=LEAF;
RX MEDLINE=20399434; PubMed=10945337;
RA Hara K., Yagi M., Koizumi N., Kusano T., Sano H.;
RT "Screening of wound-responsive genes identifies an immediate-early
RT expressed gene encoding a highly charged protein in mechanically
RT wounded tobacco plants.";
RL Plant Cell Physiol. 41:684-691(2000).
DR EMBL; AB009883; BAA95789.1; -.
SQ SEQUENCE 513 AA; 61019 MW; F9598A37A840B970 CRC64;

Query Match 22.3%; Score 327; DB 10; Length 513;
Best Local Similarity 30.2%; Pred. No. 1.3e-10;
Matches 101; Conservative 57; Mismatches 102; Indels 74; Gaps 14;

QY 2 KRNSPQIKRKDRRAK---KQSFDDNDSELEDD-----KDSKSK 38
Db 93 KKKHKTDMKEKKDKEMKDKSKHESEKDSKEIEEDDGEKKKKKKKKKKKKKK 152

QY 39 KTAQPKVEMSGSDDDDFNKLPKKAKGKAQKSNK-----KWDGSEDEDSKKIKERSRI 94
Db 153 EKDKKSIEESKEEKDDKGEKKDKQKDKKKNKEKKGKSGESEEETEEKDDKGN 212

QY 95 NSSGESS-----DESDFLQSRGKQKNQKNKPGPNIESGDDDDASPKITVAOK--- 145
Db 213 KESDEDERQTEEEENDEKGVKKDKKKKKKKKE-----KADNEKKDKS-KEETEERDDE 267

QY 146 KAEKKERERK-----KRDEE-KAKLRKLEKEEL---ETGKKDQSKQKESQKFE--- 191
Db 268 KGEKKDKKKKKKKKKKKKKDKETKDKSKVSDDEEEKDDDEGEKKDKKKKKKKKTKD 327

QY 192 -ETVSKVTVDGVPASEEKAETPAEDDNEGDGKKKKKKKKKKKKKKKKKKKKGP 250
Db 328 KKKYKSKSESE-----BEDKKETEEDDDDEGQKKEKKKKKKKKKKKKKKKKKK--- 378

QY 251 SKATVAMQEAQAKLKEEEROKREBEERIKRLE 284
Db 379 ----VKSKEES-----DEEDKQKVNEVEVATRE 403

RESULT 4
Q9VZP5 PRELIMINARY; PRT; 932 AA.
AC Q9VZP5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CG10840 PROTEIN.
GN Cif2 OR CG10840.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
```


DR Pfam: PF03144; GTP_EFTU_D2; 2.
DR PRINTS; PR00315; ELONGATNFCT.
DR PROSITE; PS00358; RIBOSOMAL_L5; UNKNOWN_1.
KW GTP-binding; Initiation factor; Protein biosynthesis.
SQ SEQUENCE 1224 AA; 135051 MW; 6C25811EB0B87002 CRC64;

Query Match 19.4%; Score 284.5; DB 10; Length 1224;
Best Local Similarity 25.6%; Pred. No. 5.1e-08;
Matches 111; Conservative 56; Mismatches 109; Indels 157; Gaps 18;

QY 1 MKNRSPQIKKKRRDRA---AKQSF--DNDSEELDKSKKTAKPK--V 45
DB 11 MGRKPSRSSRGDDEPPAASLVGAESKKKAVQSDDEKYSYNTEBEKVITGKKKSNK 70
QY 46 EMTSGSDDDDDFNKL-----PKKAGCAKSNKKW--DGSEEDENSKKIKERSR 93
D 71 KVTQKHDDDDFTAVPENGFGVKKKSKGKNGGVSFALLSGKETDDE-----122
QY 94 INSSGESDESDEFLQSRGQKK-NQNKPGPNIES-----GNEDD-----AS 136
DB 123 -----SNGKDDRPVISFTCKKNASNGKGFVAFDALGGDKDDEVDGDEEQVSPIT 177
QY 137 FKTKTVAQKAEEK-----ERKKRDE-----160
DB 178 FSGKTRSSKSKKTKTKVAPDPEEGTGMGDEESLEITFSGKKKGSTVLASLGDSVADE 237
QY 161 --KAKLRLKEKELEFGKDKDSKQESQ--RKFEETVSKVTVDGVPASEEKAET- 215
DB 238 TSQAKTPDTSKVSVEITGKTKKKKKNNKVKARTLEEEDDLKLLAELGETPAAERPASST 297
QY 216 -----PTAAEDDNEGDK-----KKDKKKKKGE-----K 239
DB 298 PEVEKVOAGPVPAPVENAGEKEKETVETAAAKKKKKKKKDKKAAKAAATSSVEAK 357
QY 240 BEKEKE-----KKGSKAT-----VRAMQALAKLKEEERQKREBEERI--- 280
DB 358 EEKQESVTEPFPKPKDKAGKAAEKKPKHVREMQBALARRQEAERKKKEEELRKE 417
QY 281 ----KRLKELESK 289
DB 418 EERRRQEELEAQ 430

RESULT 9
Q93YQ4 PRELIMINARY; PRT; 2081 AA.
AC Q93YQ4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GENOMIC DNA, CHROMOSOME 3, BAC CLONE: T19N8.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-COLUMBIA;
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RX MEDLINE-20363099; PubMed-10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety pl,
RT TAC and BAC clones."
RL DNA Res 7:217-221(2000).
DR EMBL: AP002057; BAB03174.1; -;
SQ SEQUENCE 2081 AA; 232851 MW; D3603E1F05EFF29 CRC64;

Query Match 19.3%; Score 283.5; DB 10; Length 2081;
Best Local Similarity 27.4%; Pred. No. 9e-08;
Matches 85; Conservative 60; Mismatches 98; Indels 67; Gaps 11;

QY 1 MKNRSPQIKKKRRDRA---KQSFDD-NPSEELDKSKSKKTAKPKVEMYSGSDDDDD 56
DB 906 VQKSGSVKVKYKDEKKEGKKNKDTINTSSQKQKDKKKKKSKNSNMKKKEEDKKE 965
QY 57 F--NKLPPKAKGAKQAKSNKKWDGSEEDENSKKIKERSINSSGESDESDEFLQSRGQ 114
DB 966 YVNNEL-----KQEDNKK---ETTKSENKLEENKDNKKEKSEDS----ASNRE 1011
QY 115 KKNQKNKPGPNIESGNEEDDASFKTKVAQKAEEKERKKRERKDRDEAKLRLKEKEELE 174
DB 1012 KKEYEEK-----KSKTKEAKKKEKKSKQKKREKSEERKSKKEES 1055
QY 175 TGKDKQSKQKESQKFEETVSKVTVDGVPASEEKAETPTAAEDDNEGDKKKDKKK 234
DB 1056 RDLKAKKKEETKEKESNHKSK-----KEDKKE-----HEDNKSMMKEEDKKE 1101
QY 235 KGEKEKEKEK-----KKGPSKATVKAQEAALAKLKEEERQKREBE 278
DB 1102 KKHESKSKKEEDKDKMEKLEQNSNKKKEDKNEKKKSHQHLVKKESDKKKEKNEE 1161
QY 279 RIKRLEES 288
DB 1162 K-SETKEIES 1170

RESULT 10
Q93YQ4 PRELIMINARY; PRT; 363 AA.
AC Q93YQ4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 42.4 KDA PROTEIN (FRAGMENT).
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY059831; AAL24313.1; -;
KW Hypothetical protein.
FT NON_TER 363 363
SQ SEQUENCE 363 AA; 42420 MW; DB5A4D4E487A4936 CRC64;

Query Match 19.3%; Score 283; DB 10; Length 363;
Best Local Similarity 29.2%; Pred. No. 2.2e-08;
Matches 100; Conservative 53; Mismatches 107; Indels 82; Gaps 14;

QY 1 MKNRSPQIKKKRRDRAKQSFDDN-----DSEELDKSKSKKTAKPKVEMYSGSD 52
DB 37 VKAKSIEKVRKAKRDEESGSKSKDKKKKNVDSEVKEDKDDKKDKGMVMSKHEEGH 96
QY 53 DD-----EDFNKLPKAK-----GRAQKSNKKWDGSEEDENSKRIKE 90
DB 97 GDLEVKESDVKEEHEKKGKKEKHEELEKEGKKKKKKKDESGPEKKNKADKE 156
QY 91 RSRINSSGESDESDEFLQSRGQKKNQKNKPGPNIESGNEEDDASFKTKVAQKAEEK 150


```
[1]
SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Niernman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome 1 BAC F28016 genomic sequence.";
RL Submitted (Oct-2000) to the EMBL/GenBank/DDJB databases.
DR EMBL; AC010718; AAP04442.1; -
DR InterPro: IPR000795; GTP_EFTU.
DR InterPro: IPR004161; GTP_EFTU_D2.
DR InterPro: IPR000178; IF2.
DR Pfam; PR00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFT.
KW GTP-binding; Initiation factor; Protein biosynthesis.
SQ GTP-binding; Initiation factor; Protein biosynthesis.
   QUERY MATCH          18.9%      SCORE 276.5; DB 10; LENGTH 1280;
BEST LOCAL SIMILARITY 26.0%; PRED. NO. 1.4e-07;
MATCHES 111; CONSERVATIVE 59; MISMATCHES 100; INDELS 157; GAPS 19;

QY 18 KKOSFDNDSEELDKSKKTAKPKVMYSG-----SDDDDF----- 57
DB 30 KKAQIDDDDEYSTGLUSEKSVKEEVVVITGGKKKGKNKGTQQDDDDFSDKVSAA 89
                                           |||||
QY 58 ---NKLP---KKAQGCA-----QSNKKWDCGEDE-----DN 84
DB 90 GVXDDVPETAFVAGKKKSCKGGGSVFALLDDEDEDNESGDGKDPEVISTGKKHA 149
                                           :|||:
QY 85 SKTIKE-----RSRINSSGESGDESDEFLO-----SRKGQKNKN-- 120
DB 150 SKGKKGNGNSFAASAFDALGSDDDDTEVEHDEEESPITFGKKKKSKSKKNTINSFT 209
                                           |||:::||:
QY 121 -----KPGNIENSGED-----DDAEFKITVAQKAEAKE-----REKRDEEK 161
DB 210 ADLLDEEGTDASNRSRDENTIEDSPSTFGSKSKSKKGGSVLASVGDSVADETK 269
                                           |||::||:
QY 162 AKLRKLKEKELETKG-KDOSKOKESORKPEETEIVKSKVTVDTCGVIPAS-----EE 211
DB 270 TS--DTANVEVEVGSKKKKKNNKSGRTVQEEDLDLKLAALGETPAERPASPTVEE 327
                                           |||::||:
QY 212 KAETP-----TAAEDDNEGD-----KKKKDKKKKKG-----REEKEKE 245
DB 328 KAAQPEPVAPVENAGEKEGBEETAARKKKKKKEKEKAAAAAATSVEVEKEEQEE 387
                                           |||::||:
QY 246 -----KKGPSKAT-----VKAMQPALAKLEEERQRKEEBERI-----KR 282
DB 388 SVTEPLQPKKKDAKGAEAEXKPIKHVRMOPALARQEAEREKKEEEKELURKEERRR 447
                                           |||::||:
QY 283 LEELESK 289
DB 448 QEELEAQ 454
                                           ||||:

RESULT 14
QSYP6 PRELIMINARY; PRT; 1166 AA.
ID QSYP6
AC QSYP6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DY 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE F9H16.10 PROTEIN.
DE F9H16.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX eurosids II; Brassicales; Brassicaceae; Arabidopsia.
XC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
```

SQ SEQUENCE 791 AA; 89672 MW; DA40A0FFA99E3E39 CRC64;

Query Match 18.2%; Score 267.5; DB 13; Length 791;
Best Local Similarity 24.1%; Pred. No. 2.9e-07;
Matches 70; Conservative 84; Mismatches 118; Indels 19; Gaps 6;
QY 14 DRAKQSFDDNDSEEL-----EDKSKS---KKTAKPKVEMYSGSDDDDDFNKLPKKA 64
DB 436 ERKKAASEEEGEEERADSLSEEEGDSKSDAFESDAESETE-GECAEEEQSSTEKEGES 494
QY 65 KGKAQSNKKWDGSEDEDNKKIKERSINSGSGSDSEDFLQS--RKQKKKNOKNKP 122
DB 495 NEEEQSTKEEESNEEEQSTKEEESNEEEQSTKEEESNEEEQSTKEEESNEEEQSS 554
QY 123 GPNIESGNEDDDAFQKTKVAQKKAQKKAERERKRDDEKAKLRKLKEKELETGKKKQSK 182
DB 555 TEKEESNEEEQSTKE---ESNEEEEEESESENEEEEDQKDEEEETGEEDEEE 611
QY 183 QKESQRKFEETVKSQVTDGTVIPASEEKAETPTAAEDDNEGDKKKKDKKKK---GE 238
DB 612 EEESEQNEEEAEEDGTDVQEEEEEETETEKEEEDDEEETEKEEEDDEEETDRVEE 671
QY 239 KEEKEKKKGPQKATVKAMQALAKLKEEERQKREBEERIKRLEELSK 289
DB 672 EDAEENEEEEESENEEEDSDGEEDEEDSDSEDESESESEEEERGE 722

Search completed: July 19, 2002, 14:18:09
Job time: 221 sec